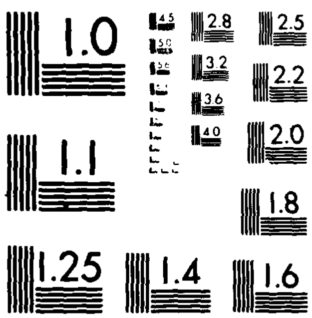
AD-A108 681 AIR FORCE GEOPHYSICS LAB HANSCOM AFB MA F/G 4/2
AN AUTOMATED LOW CLOUD PREDICTION SYSTEM.(U)
JUL 81 E B GEISLER
UNCLASSIFIED AFGL-TR-81-0191 NL

1 of 1

END
DATE
FILMED
82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

(12)

AFGL-TR-81-0191

ENVIRONMENTAL RESEARCH PAPERS, NO. 746

AD A108681

An Automated Low Cloud Prediction System

EDWARD B. GEISLER, Capt, USAF

7 July 1981

Approved for public release; distribution unlimited.

METEOROLOGY DIVISION PROJECT 6670
AIR FORCE GEOPHYSICS LABORATORY
HANSCOM AFB, MASSACHUSETTS 01731

AIR FORCE SYSTEMS COMMAND, USAF

This report has been reviewed by the ESD Information Office (OI) and is releasable to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

John Howard

Chief Scientist

Qualified requestors may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFGL-TR-81-0191	2. GOVT ACCESSION NO. AD-A108/81	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) AN AUTOMATED LOW CLOUD PREDICTION SYSTEM		5. TYPE OF REPORT & PERIOD COVERED Scientific. Final
		6. PERFORMING ORG. REPORT NUMBER ERP No. 746
7. AUTHOR(s) Edward B. Geisler, Capt, USAF		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Geophysics Laboratory (LYU) Hanscom AFB Massachusetts 01731		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62101F 66701004
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Geophysics Laboratory (LYU) Hanscom AFB Massachusetts 01731		12. REPORT DATE 7 July 1981
		13. NUMBER OF PAGES 40
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION ST   1ENT (of   abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY   TES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Rotating beam ceilometer
Wind sensor
Forward scatter meter
Weather test facility
Probability forecast techniques
Categorical forecast techniques

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

At the Air Force Geophysics Laboratory (AFGL) Weather Test Facility (WTF) at Otis Air Force Base, Massachusetts, a network of cloud base height, visibility, and wind measuring instruments were utilized to explore techniques for the short range prediction of low cloud ceiling. AFGL has developed this system in response to the USAF Air Weather Service's requirements to modernize its basic weather support capabilities. This system allowed AFGL to evaluate the ability of statistical forecasting techniques to provide decision assistance significantly improved over the decision assistance currently

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. (Cont)

provided by climatology and persistence. The approach relies upon the use of a hierarchical clustering algorithm to transform the raw cloud base height data into an automated low cloud observation. Four prediction techniques (Regression Estimation of Event Probabilities, Equivalent Markov, climatology, and persistence) yielding probability estimates of low cloud ceiling were evaluated and comparisons made to determine their respective accuracy and reliability. In addition, thresholding techniques were used to convert probability forecasts (unit bias, maximum probability, iterative, and persistence).

Analysis of the data collected at the AFGL WTF demonstrates the accuracy and reliability of the automated low cloud prediction system. Regression estimation of event probabilities provided accurate, reliable, high resolution probability forecasts with results superior to climatology, persistence, and Equivalent Markov. Examination of the categorical forecast techniques showed that persistence provided the simplest method while yielding extremely competitive results. Forecast lengths considered were 30, 60, 120, and 180 minutes.

Preface

This work has benefitted from the help of many people without whom successful completion would not have been realized. The author is especially grateful to Mr. Donald A. Chisholm for many helpful discussions on the prediction techniques and for reviewing the report; Capt. James C. Weyman and Richard Lynch for their management of the Modular Automated Weather System (MAWS); Leo Jacobs, Ralph Hoar and Clyde Lawrence for maintaining the field test instrumentation; Joan Ward for extensive software development and data processing; William Lamkin for illustration preparation; Maureen Hampton for developing plotting software; and to Karen A. Sullivan for typing the manuscript.

Contents

Contents

Illustrations

Tables

Tables

An Automated Low Cloud Prediction System

1. INTRODUCTION

The USAF Air Weather Service (AWS) recognizes the need to modernize its weather support capabilities. The Automated Weather Distribution System (AWDS) Multi-Command Required Operational Capability (ROC-801-77) and the AWS GR 3-76 (Automated Short-Range Terminal Forecasts) identify the need to provide responsive environmental support to local decision makers. This support, in the form of short-range forecasts, must provide a statistically significant improvement over decision assistance provided by persistence. In response to these USAF requirements, the Air Force Geophysics Laboratory (AFGL) has developed forecast schemes which provide probability estimates for several critical weather elements: radiation fog,[1] slant visual range,[2] and visibility.[3]

The objective of this study is to develop automated procedures for probability and categorical forecasts of low cloud ceiling, defined here to be less than 6000 ft. The approach relies upon the automated cloud observing system developed by AFGL

(Received for publication 22 June 1981)

1. Tahnk, W.R. (1975) Objective Prediction of Fine Scale Variations in Radiation Fog Intensity, AFCRL-TR-75-0269, AD A014774.
2. Geisler, E.B. (1979) Development and Evaluation of a Tower Slant Visual Range System. AFGL-TR-79-0209, AD A082384.
3. Hering, W.S., and Quick, W.L. (1974) Hanscom visibility forecasting experiment, Proc. of 5th Conference on Weather Forecasting and Analysis, pp 224-227.

to provide the necessary data bases to generate and verify forecast equations. Climatology, persistence, Equivalent Markov, and regression estimation of event probabilities (REEP) prediction algorithms were applied to generate probability forecasts. Categorical forecasts were retrieved from the probability forecasts by four techniques: maximum probability, persistence, unit bias, and iterative. The probability and categorical forecast techniques were evaluated and comparisons made to determine their respective accuracy and reliability. Forecast lengths examined in this study are 30, 60, 120, and 180 min.

2. PROBABILITY FORECAST TECHNIQUES

2.1 Regression Estimation of Event Probabilities (REEP)

REEP[4] is analogous to a forward screening regression technique where a subset of predictors is selected from a large set of possible predictors. The relationship between the predictand and the selected predictors is a linear function whose coefficients are determined by least squares. REEP calculates probabilities of being within categories. These within-category probabilities can be easily converted into exceedance probabilities. REEP has been successfully used in other forecasting experiments conducted at AFGL by Tahnk[1] and Geisler.[2]

REEP requires each raw variable be broken down into a set of two or more mutually exclusive and exhaustive categories called dummy variables. The dummy variable categories are assigned a value of 1 if the continuous variable falls within the range of the category limits, or 0 if it is outside the range. Therefore, for any one raw variable broken down into n dummy variables, one dummy variable will be assigned a value of 1 and all of the others (n-1) will be assigned a value of 0.

The result of the least squares technique is a set of equations, one equation for each predictand category (i).

$$P_i = A_{0i} + \sum_{j=1}^{k} A_{ij} X_j, \quad i = 1, 2, \ldots, 5 \tag{1}$$

where k is the number of dummy predictors (X_j) selected by a screening process, and the A's the regression constants and coefficients.

REEP is formulated to insure internal consistence among predictand categories such that the sum of probabilities totals unity. This is insured by using the same predictor categories for each predictand equation. Each predictand equation yields a probability of occurrence (P_i) for "its" category.

4. Miller, R.G. (1964) Regression Estimation of Event Probabilities, Tech. Rpt. 7411-121, The Travelers Research Center, Inc., Hartford, Connecticut

2.2 Equivalent Markov Based on REEP

Miller[5] proposed a prediction method yielding probabilistic forecasts comparable to the classical Markov process but without the necessity of utilizing the Markov transition matrix $\underset{\sim}{M}$ explicitly. Whiton[6] and Geisler[2] found this technique to yield results comparable to the classical Markov process but much easier to develop and apply to practical forecasting problems.

The REEP equations provide a conditional probability matrix $\underset{\sim}{P}$ equivalent to $\underset{\sim}{M}$. One assumes the categories (predictor and predictand) used in REEP form a finite Markov chain (X_i), $i = 1,2,\ldots,5$. Hence, $\underset{\sim}{P}$ is a square matrix.

The stochastic process (X_i) has the Markovian property

$$P(X_{t+1} = i \mid X_t = j) = P_{ij}; \qquad i, j = 1, 2\ldots, 5 \; . \tag{2}$$

The existence of one-step transition probabilities P_{ij} implies that for each i, j and n

$$P(X_{t+n} = i \mid X_t = j) = P(X_n = i \mid X_o = j) = P_{ij}{}^{(n)} \; . \tag{3}$$

$P_{ij}{}^{(n)}$ is the conditional probability the random variable X, starting in state j, will be in state i after exactly n steps (time units) and is called the n-step transition probability.

More generally, the matrix of n-step transition probabilities can be obtained from the expression:

$$\underset{\sim}{P}^{(n)} = \underset{\sim}{P}.\underset{\sim}{P}\ldots\underset{\sim}{P} = \underset{\sim}{P}^{n} = \underset{\sim}{P}\cdot\underset{\sim}{P}^{n-1} = \underset{\sim}{P}^{n-1}\underset{\sim}{P} \; . \tag{4}$$

Thus, the n-step transition probability matrix can be determined by computing the nth power of the one-step transition matrix $\underset{\sim}{P}$. In this application of the Equivalent Markov model, a 30-min one-step transition matrix was used to generate predictions from 30 to 180 min (see Appendix B). For example, to generate the 120-min forecast probabilities, one must first raise $\underset{\sim}{P}$ to a power of 4 and then pre-multiply the resultant transition matrix by $\underset{\sim}{O}$, the observation matrix:

$$\underset{\sim}{O} = (O_1, O_2, \ldots, O_k) \tag{5}$$

5. Miller, R.G. (1968) A stochastic model for real-time on-demand weather predictions, Proc. 1st Statistical Meteorological Conference, Am. Meteor. Soc. pp 48-51.

6. Whiton, R.W. (1977) Selected Topics in Statistical Meteorology, AWS-TR-77-273, pp 7-1 to 7-45.

where

k = predictor category number

$$O_k \left\{ \begin{array}{l} \text{zero, if category is not observed} \\ \\ \text{one, if category is observed} \end{array} \right\}$$

2.3 Climatology

Climatology forecasts were computed using the Wind Stratified Conditional Climatology (WSCC) and Revised Uniform Summary of Surface Weather Observations (RUSSWO) climatology tables. These tables were based on approximately 20 years of hourly human observations taken at Otis AFB, located on Cape Cod, Massachusetts.

These tables have common characteristics but have some significant differences. Both tables are stratified so one can quickly retrieve the frequency of occurrence for any event (for example, ceiling less than 200 ft), conditioned on month and time of day. However, the WSCC table is further divided to condition the event on wind direction and cloud ceiling. This allows one to retrieve from the WSCC table, forecasts for the frequency of occurrence of cloud ceiling for periods of 1, 2, and 3 h. This is unlike RUSSWO which provides only climatological frequencies for a particular time of day and month.

In our use of the WSCC table, we chose to condition the predictand event, cloud ceiling, on three variables: month, time of day, and initial cloud ceiling category. If we had elected to also condition the event on wind direction, some of the frequency of occurrence values obtained from the WSCC table would have been statistically unstable, because they were based upon too few observations.

In generating climatological forecasts only one table was used for each forecast made. If a ceiling condition was reported, the WSCC table was used. A 30-min climatology forecast was not made because the WSCC table only provides hourly forecasts. If a no ceiling condition was observed, then the RUSSWO table was used.

2.4 Persistence

The model used as a control technique in this experiment was persistence. Here, one takes the initial cloud ceiling category, assigns a probability of 1.0 to it and 0.0 to the remaining categories and assumes that conditions will not change during the period of the forecast.

3. INSTRUMENTATION AND DATA SETS

3.1 Instrument Configuration

Measurements of cloud base height (AN/GMQ-13 Rotating Beam Ceilometer: RBC), wind direction (Climatronics Mark I Wind Sensor), and visibility (EG&G Model 207 Forward Scatter Meter: FSM) were obtained at the AFGL Weather Test Facility (WTF) located at Otis AFB, Massachusetts. The FSM and Climatronics Wind Sensor have been used in several experiments[7] carried out by AFGL during the past several years. Figure 1 shows the configuration of the surface-based instruments used in the study.

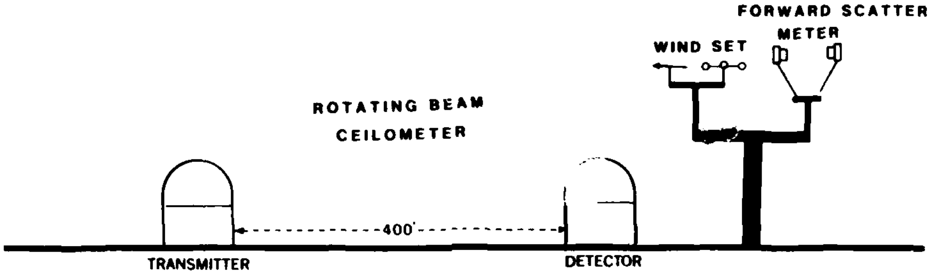


Figure 1. Instrument Configuration

The RBC[8] was used to acquire cloud height data in these tests. The RBC is the standard cloud height measuring device used by the USAF Air Weather Service, the National Weather Service, and internationally. It consists of a two-lamp projector system and a receiver which is normally set 400 ft from the projector with its field of view vertical and coplanar with the rotating projector beams. The intersection volume moves upward as each projector beam scans to the vertical. When the volume intersects a cloud, backscatter of the projector beam by particles in the cloud is detected by the receiver and displayed on the indicator as a height vs intensity depiction. The projector angle at which the maximum backscatter return occurs represents the cloud height. Cloud heights in our tests were limited to measurements up to 6000 ft because of known sensor and trigonometric limitations and after an assessment of basic RBC data.

---

7. Chisholm, D. A., Lynch, R. H., Weyman, J. C., and Geisler, E. B. (1980) A Demonstration of the Modular Automated Weather System (MAWS), AFGL-TR-80-0087, AD A087070.

8. Weyman, J. C., and Lynch, R. H. (1981) A Digital Processing and Display System for the Rotating Beam Ceilometer (AN/GMQ-13), AFGL-TR-81-0027.

Measurements of visibility were obtained from the FSM in extinction coefficient units. Data were collected at a rate of 5 observations per min and were subsequently processed to yield a 1-min average of extinction coefficient. One can use Koschmeider's Law (for daytime conditions) and Allard's Law (for nighttime conditions) to explicitly relate visibility to extinction coefficient. Figure 2 shows representative values of day and night visibility (in meters) for a given value of extinction coefficient (in units of km^{-1}).

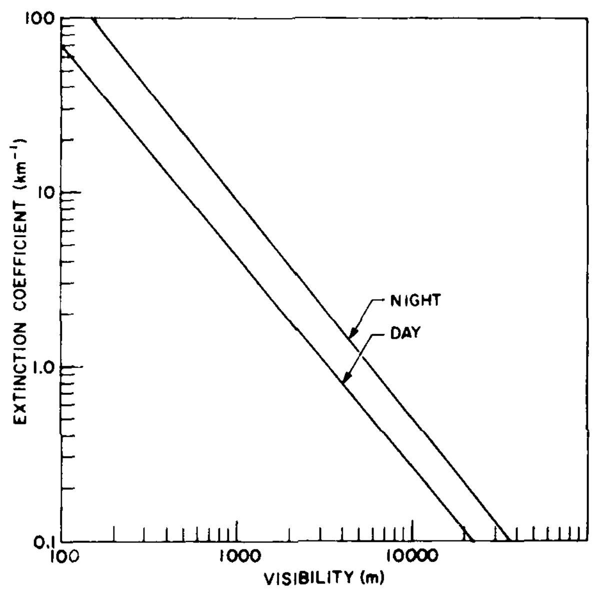


Figure 2. Relation of Extinction Coefficient to Visibility

The Climatronics cup-and-vane wind sensor is lightweight, has low power consumption and a low start-up threshold. (A wind speed of 0.22 m/sec turns the cups and 0.11 m/sec moves the wind direction vane.) It also responds quickly and is very accurate (for wind speed, ± 1 percent or ± 0.7 m/sec, whichever is greater, and for direction ± 2.5°). A non-contacting wind-direction transducer and a chopped solid-state light source sense direction and speed. Measurements were collected at a rate of 1 observation per sec and processed to yield a 1-min average of wind direction and speed.

3.2 Data Sets

Data were collected from March to December 1980. Wind direction, visibility and cloud base height measurements were placed on magnetic tape and subsequently processed on the CDC 6600 system. This processing transformed the raw cloud base height values into an objectively determined low cloud observation by applying the hierarchical clustering technique. The hierarchical clustering technique has been shown to yield stable, reliable, and representative cloud observations.[9, 10] It is these automated cloud observations which made up the population from which episodes were selected for this study. In all, 43 episodes were chosen. Episodes were selected if they satisfied three criteria: (1) episode duration was at least 3 h, (2) clouds were reported below 6000 ft, and (3) a ceiling was reported at some time during the episode. Six of the episodes, containing approximately 7000 observations, were randomly chosen and set aside as an independent data set. The remaining 37 episodes, comprised of approximately 20,000 observations, made up the dependent data set.

Four predictand categories were selected to coincide with the first four cloud ceiling categories of the WSCC and the remaining events in the sample comprised a fifth category. Table 1 lists the category limits.

Table 1. Predictand Category Limits

Predictand Category	Limits (ft)
1	Ceiling < 200
2	200 ≤ Ceiling < 500
3	500 ≤ Ceiling < 1000
4	1000 ≤ Ceiling < 3000
5	3000 ≤ Ceiling < 6000 or No Ceiling

Table 2 summarizes the relative frequency distribution for the developmental and independent data sets. Note that the developmental data frequency distribution represents a randomly selected sub-sample taken from the dependent data set. The limit on sample size of 3000 is imposed by the REEP computer programs.

---

9. Geisler, E. B., and Chisholm, D. A. (1980) An Automated Cloud Observing System (ACOS) *AFGL-TR-81-002, AD A100266.*

10. Duda, R. O., Mancuso, R. L., and Paskert, P. F. (1971) Analysis of Techniques for Describing the State of the Sky Through Automation, Report No. FAA-RD-71-52, prepared for Department of Transportation, Washington, D. C.

Table 2. Frequency Distribution of Independent and Developmental Data by Predictand Categories

Predictand Category	Developmental Observations	Independent Observations
1	145	120
2	547	1211
3	917	2276
4	674	1547
5	717	1873
Total	3000	7027

Three raw variables were selected to constitute the predictor set: Low cloud observation, visibility, and wind direction. Category limits were assigned to visibility and wind direction based on their relative frequency distributions in the dependent data set. The resultant categories for the predictors are listed in Table 3. The predictor categories were used as input to each of the probability prediction techniques.

Table 3. Predictor Categories

Predictor Category	Type (Units)	Limits
1	Cloud Height (ft)	Ceiling < 200
2	Cloud Height (ft)	200 < Ceiling < 500
3	Cloud Height (ft)	500 ≤ Ceiling < 1000
4	Cloud Height (ft)	1000 ≤ Ceiling < 3000
5	Cloud Height (ft)	3000 ≤ Ceiling ≤ 6000 Or No Ceiling
	Extinction Coefficient	
6	EC (km^{-1})	EC ≤ 0.4
7	EC (km^{-1})	0.4 < EC < 2.0
8	EC (km^{-1})	2.0 < EC ≤ 25.0
9	EC (km^{-1})	25.0 < EC
10	Wind Direction (deg)	WD ≤ 320 or WD < 100
11	Wind Direction (deg)	100 ≤ WD < 160
12	Wind Direction (deg)	160 < WD < 230
13	Wind Direction (deg)	230 ≤ WD < 320

Three variations of possible predictors were evaluated for REEP. This allowed a determination of the additive contribution of visibility and wind direction information to prediction accuracy beyond that attained using only the initial cloud observation as a predictor. Hence, three REEP equation sets were generated from a randomly selected sub-sample of 3000 observations taken from the dependent data set. The first set, REEP5, used only the initial cloud observation. The second set, REEP.01, used selected predictors which reduced the predictand variance by at least 1 percent. The final set, REEP11, used all the predictors selected by least squares with no restriction on the amount of variance reduced. Appendix A lists the REEP equation sets.

4. CATEGORICAL FORECAST TECHNIQUES

There are a variety of methods by which a categorical forecast can be made. In this experiment, we chose to compare four methods: maximum probability, iterative, unit bias, and persistence. The first three methods are thresholding processes which transform a probability forecast into a categorical forecast. NOAA/TDL[11] has used thresholding methods with excellent results. Since the thresholding approach requires probability estimates of the events forecasted, a REEP equation set was generated from the developmental sample. The categories or events were (1) ceiling less than 1000 ft, and (2) either ceiling greater than or equal to 1000 ft or no ceiling. In general, if the REEP probability exceeds the threshold probability (p^*) for the first category, then the first category is chosen to be the categorical forecast. Otherwise, the second category is selected. The specific methods of assigning the threshold probability are discussed below. Appendix C lists the REEP equation set used here. Categorical forecasts were made for all four forecast intervals (30, 60, 120, and 180 min).

4.1 Maximum Probability

The maximum probability (Max Prob) model selects the category with the largest probability as the categorical forecast.[12] This is equivalent to a thresholding technique with $p^* = 0.5$. It is perhaps the most commonly used because of its simplicity. REEP and climatology probabilities were used by the model. REEP (Max Prob R) is discussed in Section 2.1 and climatology (Max Prob c) in Section 2.3. The use of both methods to convert from probability to categorical forecasts enables us to determine any difference in skill between them in making the conversion.

---

11. Miller, R.G., and Best, D.L. (1978) A model for converting probability forecasts to categorical forecasts, NOAA Tech. Memo, TDL Office Note 78-14.

12. Miller, R.G., and Best, D.L. (1979) A model for converting probability forecasts to categorical forecasts, Preprints, Sixth Conf. on Probability and Statistics in Atmospheric Sciences, Am. Meteor. Soc., pp 98-102.

4.2 Iterative

The iterative technique is an empirical one. On successive passes through the dependent data set, the descriptive statistics bias, threat score, and percent correct (see Section 5.2) were computed by comparing the actual probability forecasts against incremented threshold probabilities. The selected threshold probability is then subjectively determined based mainly on the bias statistics. Table 4 lists the threshold probabilities selected for each forecast interval, for the predictand category ceiling < 1000 feet.

Table 4. Threshold Probabilities for Ceiling Less Than 1000 ft Selected From Iterative Method

Forecast Length (min)	Threshold Probability of Ceiling < 1000 ft
30	0.72
60	0.69
120	0.58
180	0.64

4.3 Unit Bias

The unit bias model seeks to establish thresholds which will result in a categorical bias statistic of one (unity) for the developmental sample. Threshold probabilities are calculated by the unit bias model; however, they are determined differently from the iterative technique. Here, p^* is a function of the climatology (c) of the predictand event, ceiling less than 1000 ft, and the multiple correlation coefficient (R) between the predictand and predictors in the forecast equation. This information permits a fast and simple calculation of p^*, compared to the iterative technique:

$$p^* = R\,(0.5 - c) + c \qquad (6)$$

Table 5 lists the (R, c) pairs for each forecast length.

4.4 Persistence

The control technique was persistence. Whatever the initial cloud category was at time zero, the categorical forecast was to keep it in that category for all forecast lengths.

Table 5. Unit Bias Correlation and Climatology (R, c) Pairs

Forecast Length (min)	R	c
30	0.849	0.452
60	0.778	0.455
120	0.679	0.465
180	0.594	0.466

5. EVALUATION CRITERIA

5.1 Probability Forecasts

Probability estimates of a weather event must be accurate and reliable to be useful. Over a period of time, the event should actually occur with the frequency represented by the probability forecast. In addition, resolution in the probabilities is desirable; that is, they should be as close to 0.0 or to 1.0 as possible when the event does not occur or does occur, respectively.

The accuracy of the probability forecasts was determined by the Brier Probability (or p) Score[13] which measures the mean-squared probability errors. It is determined from Eq. (7).

$$p = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{k} (P_{ij} - O_{ij})^2 \qquad (7)$$

where

N number of forecasts,
P_{ij} probability estimate for jth predictand category,
O_{ij} observed value (0, 1) for jth predictand category,
k number of predictand categories.

The range of values for p depends on the number of predictand categories. For two or more predictand categories, p ranges from 0.0 to 2.0. For one predictand (ceiling less than 1000 ft) p ranges from 0.0 to 1.0. A perfect probability forecast has a Brier Score of 0.0 and the worst probability forecast has the maximum value in the range.

Reliability of the probability forecasts can be depicted by a simple graphing technique which allows us to compare the predicted probability distribution to the observed frequency. Figure 3 is an example of a reliability graph in which eleven

13. Brier, G.W. (1950) Verification of forecasts expressed in terms of probability, Mon. Wea. Rev., 78:1-3.

probability categories and observation frequencies are used. The diagonal line denotes perfect reliability where, for example, 40 percent of the cases in which the probability forecasts ranged from 0.35 to 0.45 resulted in the forecasted event actually occurring. The bar graph to the right depicts the proportion of probability forecasts in each of the eleven categories. The categories used are

[(0,.05), (.05,.15), (.15,.25), ..., (.85,.95), (.95,1.0)] .

In this illustration the distribution of probability forecast reflects fairly poor resolution because the majority of the probabilities are "mid-range" values (20-80 percent).

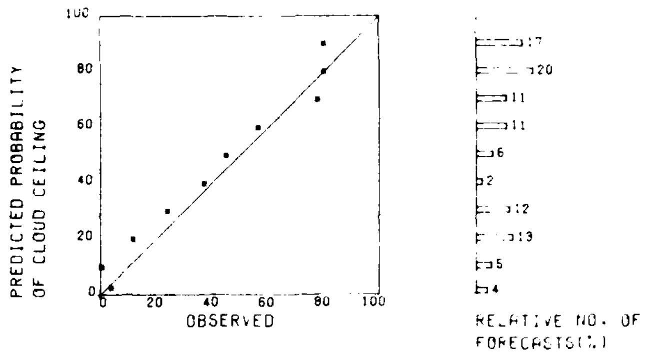


Figure 3. Reliability Graph Example

5.2 Categorical Forecasts

Categorical forecasts were made for one of two events: (1) ceiling less than 1000 ft, or, (2) ceiling greater or equal to 1000 ft, which includes no ceiling cases.

Categorical forecast skill was measured using a standard set of descriptive statistics: bias, threat score, and percent correct. Bias is defined here as the fraction (F/O) where F and O are the number of forecasts and observations of the same category, respectively. Unit bias is attained when F and O are equal. Threat score is defined as $[H_1/(F+O-H_1)]$ where H_1 is the number of correct forecasts of a particular category (for example, ceiling less than 1000 ft), while F and O are as defined above. Percent correct is defined as $(H_1+H_2)/TOT$ where (H_1+H_2) is the total number of correct forecasts and TOT is the sample size. The data in Table 6 can be used to describe the method of calculating the statistics used here.

Table 6. Prediction and Observation Data Used to Calculate Skill Measures for Categorical Forecasts

Observed	Forecast		
	Ceiling < 1000 ft	Ceiling > 1000 ft or No Ceiling	Total
Ceiling < 1000 ft	2665	227	2892
Ceiling > 1000 ft or No Ceiling	216	2589	2805
Total	2881	2816	5697

For the category ceiling < 1000 ft, descriptive statistics are

$$\text{Bias} = (2892/2881) = 1.004$$

$$\text{Threat Score} = (2665)/(2892 + 2881 - 2665) = 0.857$$

$$\text{Percent Correct} = 100(2665 + 2589)/5697 = 92.2$$

6. COMPARATIVE ANALYSIS

The probability and categorical forecast techniques were applied to each episode in the independent data set. We combined the results of each episode to determine the overall accuracy and reliability of each forecast technique.

6.1 Probability Forecasts

In our examination of forecast techniques, we wanted to demonstrate the feasibility of applying finely tuned statistical methods rather than standard routines (for example, climatology and persistence). Accordingly, we designed two scenarios. First, we wanted to determine the additive contribution of visibility and wind direction information (REEP11 and REEP.01) to prediction accuracy beyond using the initial cloud observation (REEP5 and Equivalent Markov) as a predictor. The five predictand categories shown in Table 1 were used in this first scenario. Second, we wanted to compare the REEP and climatology techniques. Since the climatology tables did not provide all five predictand categories listed in Table 1, a one-category predictand (ceiling less than 1000 ft) was selected as a means to evaluate these techniques. In both scenarios, persistence was used as the control technique. The techniques were compared by use of the Brier Score (p) statistic which calculates the mean squared probability error and by reliability graphs.

For the five-predictand scenario, verification results of the p-score in Figure 4, show that REEP11 yields a slight improvement over REEP.01, REEP5 and Equivalent Markov as well as a significant improvement over persistence. The 5 percent maximum improvement of REEP 11 over Equivalent Markov, when both are compared with persistence (Table 7), suggests that the additive contribution of visibility and wind direction to prediction accuracy is small. This, in turn, suggests one could use Equivalent Markov in lieu of REEP11 with a slight decrease in accuracy. A single equation set could then be used for all forecast intervals. However, an additional penalty is incurred because of the poor resolution of the Equivalent Markov model. This can be seen in Figures 5 and 6 for category 3, ceiling between 500 and 1000 ft. Note that the reliability of both models is very good. However, Equivalent Markov generates no forecasts in the 0-15 percent range while 23 percent of the REEP11 forecasts lie in this range. This poor resolution of Equivalent Markov holds for all predictand categories, and therefore, results in part, in higher (less accurate) p-scores. We therefore conclude that REEP11 is the preferred forecast technique.

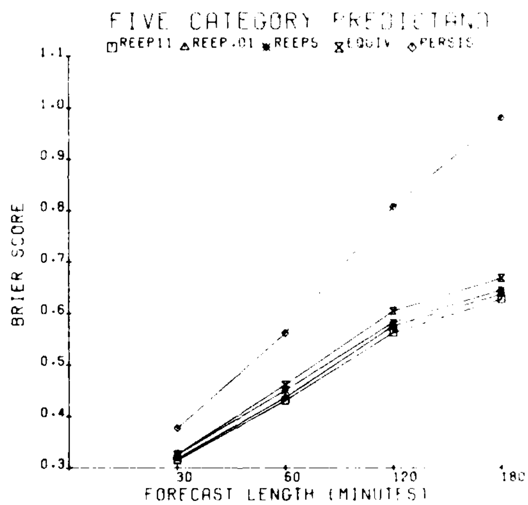


Figure 4. Brier Scores for All Forecasts Using a Five-Category Predictand

Table 7. Percent Improvement Over Persistence in Five-Category Brier Score

Prediction Technique	Forecast Length (min) 30	60	120	180
REEP11	16.5	23.3	30.3	35.9
REEP.01	15.6	22.2	28.7	34.9
REEP5	13.5	19.8	27.9	34.2
Equivalent Markov	13.5	17.8	25.7	31.8

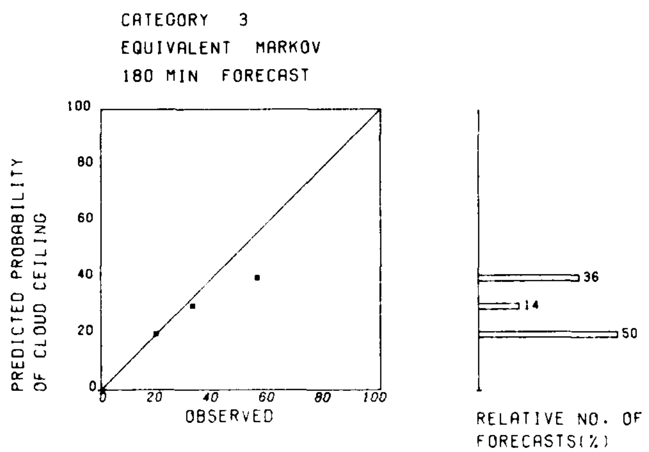


Figure 5. Reliability Graph of Forecasts From the Equivalent Markov Model for 180 min, Ceiling 500-1000 ft

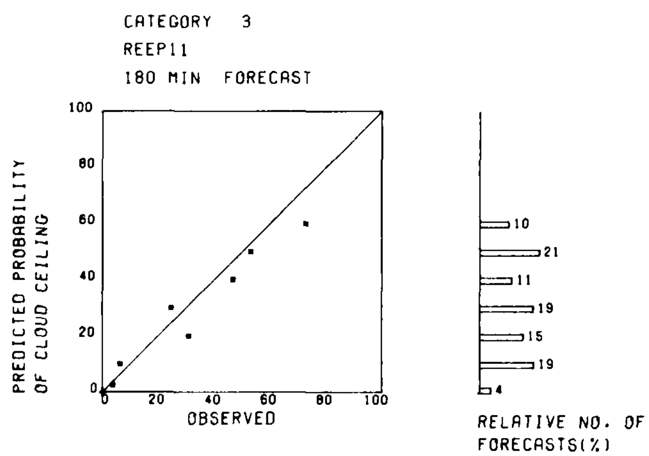


Figure 6. Reliability Graph of Forecasts From the REEP11 Model for 180 min, Ceiling 500-1000 ft

For the one-predictand scenario, categorical probabilities were converted to exceedance probabilities. This was simply done after the (P_1, P_2, P_3, P_4, P_5) were known. For example, the probability a ceiling will be below 1000 ft (P_3 or less) given (0.2, 0.5, 0.1, 0.15, 0.05) is

$$P(\text{ceiling} < 1000 \text{ ft}) = P_1 + P_2 + P_3 .$$

Hence

$$P(\text{ceiling} < 1000 \text{ ft}) = 0.8 .$$

Verification results of the p-score are shown in Figure 7. They show that REEP11 yields a significant improvement over Equivalent Markov, climatology, and persistence. Also, REEP11 yields marginal improvement over REEP.01 and REEP5. The improvement of REEP11 over Equivalent Markov is as much as 10 percentage points (see Table 8), with the underlying reason the poor forecast resolution of the Equivalent Markov model. Figures 8 and 9 show that while REEP11 makes approximately 46 percent of its forecasts in the (0 to 5) and (85 to 100) percent ranges, Equivalent Markov does not make any. Figure 10 shows that the poor performance of climatology is due to the systematic bias to underforecast and its relatively poorer resolution. Therefore, as it is in the five-predictand case, REEP11 is the most accurate and reliable forecast technique for a single predictand.

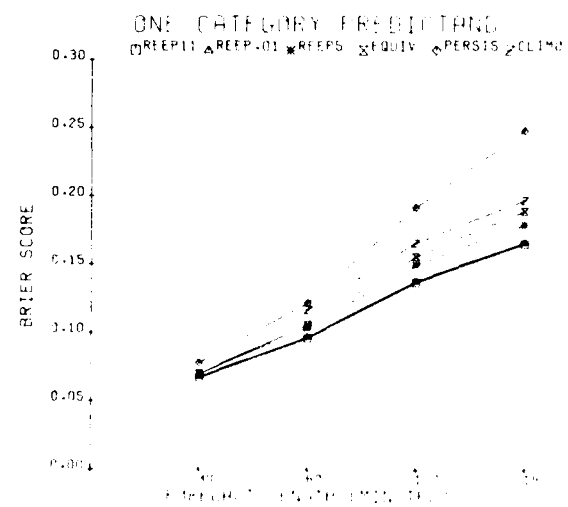


Figure 7. Brier Scores for All Forecasts Using a One-Category Predictand

Table 8. Percent Improvement Over Persistence in One-Category Brier Score

Prediction Technique	Forecast Length (min) 30	60	120	180
REEP11	14.1	21.3	28.6	33.5
REEP.01	12.8	20.5	28.1	33.1
REEP5	10.3	14.8	21.9	27.8
Equivalent Markov	10.3	13.1	18.8	23.8
Climatology	----	4.1	13.6	20.6

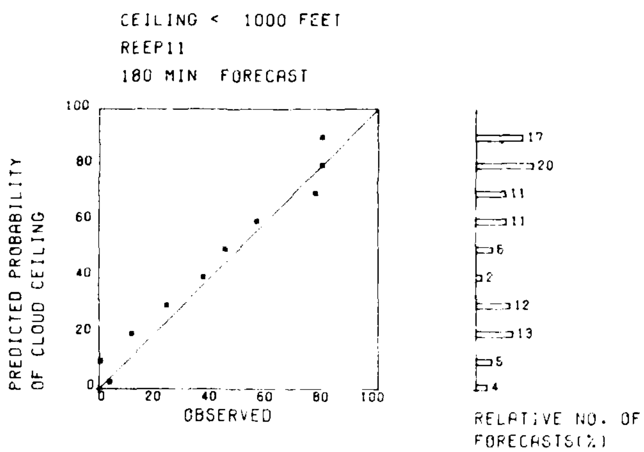


Figure 8. Reliability Graph of Forecasts From the REEP11 Model for 180 min, Ceiling < 1000 ft

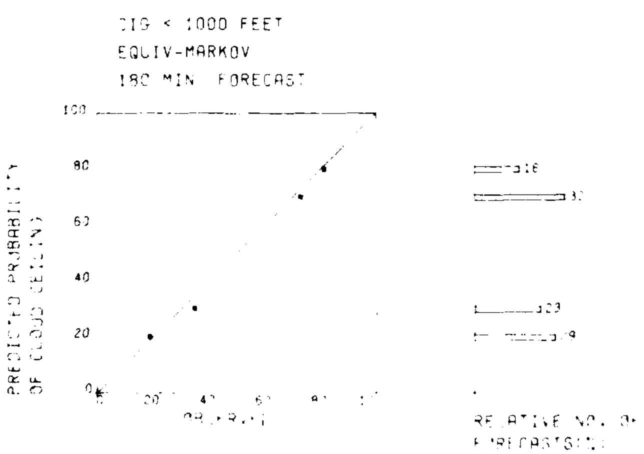


Figure 9. Reliability Graph of Forecasts From the Equivalent Markov Model for 180 min, Ceiling < 1000 ft

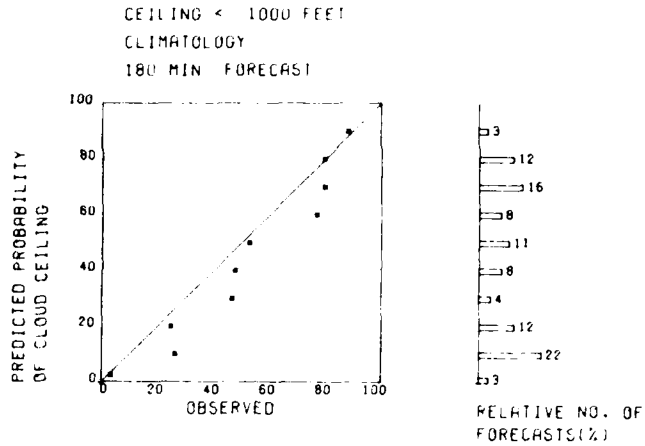


Figure 10. Reliability Graph of Forecasts From Climatology for 180 min, Ceiling < 1000 ft

6.2 Categorical Forecasts

The two-category predictand enabled us to compare various techniques for converting a probability estimate into a categorical forecast. Three significant results can be drawn from Tables 9 through 12, categorical forecast statistical summaries. One objective was to compare converting REEP and climatology probabilities into categorical forecasts using the maximum probability model, Max Prob R and Max Prob c, respectively. Here, REEP consistently yielded marginally better results than climatology. Also, REEP systematically overcast the event while climatology underforecast it. The second objective was to assess the relative skill of the iterative method and the comparatively simple unit bias method. The unit bias method yields results superior to the iterative method coinciding with findings made by TDL.[12] In effect, one can simply calculate p^* by having the (R, c) pair rather than having to further utilize the developmental sample. Finally and most important, persistence forecast results are extremely competitive with the other methods. Note that persistence forecast the event as often as it occurs and yielded similar threat and percent correct statistics.

Table 9. Descriptive Statistics for 30-min Categorical Forecasts

Method	Bias	Threat Score	Percent Correct
Max Prob R	1.004	0.857	90.6
Persistence	1.004	0.857	92.2
Iterative	0.908	0.799	84.2
Unit Bias	1.004	0.857	90.6

Table 10. Descriptive Statistics for 60-min Categorical Forecasts

Method	Bias	Threat Score	Percent Correct
Max Prob R	1.002	0.786	87.8
Max Prob c	0.975	0.766	86.7
Persistence	1.002	0.786	87.8
Iterative	0.900	0.731	85.9
Unit Bias	1.002	0.786	87.8

Table 11. Descriptive Statistics for 120-min Categorical Forecasts

Method	Bias	Threat Score	Percent Correct
Max Prob R	1.090	0.696	79.9
Max Prob c	0.966	0.642	78.0
Persistence	0.967	0.665	80.7
Iterative	0.927	0.675	80.6
Unit Bias	1.209	0.696	79.7

Table 12. Descriptive Statistics for 180-min Categorical Forecasts

Method	Bias	Threat Score	Percent Correct
Max Prob R	1.062	0.627	75.6
Max Prob c	0.926	0.567	72.5
Persistence	1.003	0.613	75.2
Iterative	0.900	0.595	75.2
Unit Bias	1.156	0.661	77.3

7. SUMMARY AND CONCLUSIONS

Analysis of the data collected at the AFGL Weather Test Facility demonstrated the accuracy and reliability of finely tuned statistical methods for low cloud ceiling prediction. In an examination of probability forecasts, REEP was found to yield significantly better results on independent data than did the Equivalent Markov technique, climatology, and persistence. In addition, visibility and wind direction information yielded marginal improvement to the accuracy and reliability of the REEP probability forecasts beyond that already attained using only the initial cloud observation as a predictor. Specifically, REEP11 had the greatest accuracy and superior forecast resolution.

In making categorical forecasts, persistence yielded extremely competitive results on independent data when compared with the thresholding methods.

It is concluded that these techniques provide timely, cost-effective decision aids to the local decision maker. Therefore, they satisfy, in part, AWS requirements in modernizing its basic weather support capabilities. The development and examination of these forecast techniques is a part of AFGL's continuing investigation of automated airfield weather observation and prediction systems.

References

1. Tahnk, W.R. (1975) Objective Prediction of Fine Scale Variations in Radiation Fog Intensity, AFCRL-TR-75-0269, AD A014774.
2. Geisler, E.B. (1979) Development and Evaluation of a Tower Slant Visual Range System, AFGL-TR-79-0209, AD A082334.
3. Hering, W.S., and Quick, W.L. (1974) Hanscom visibility forecasting experiment, Proc. of 5th Conference on Weather Forecasting and Analysis, pp 224-227.
4. Miller, R.G. (1964) Regression Estimation of Event Probabilities, Tech. Rpt. 7411-121, The Travelers Research Center, Inc., Hartford, Connecticut.
5. Miller, R.G. (1968) A stochastic model for real-time on-demand weather predictions, Proc. 1st Statistical Meteorological Conference, Am. Meteorol. Soc., pp 48-51.
6. Whiton, R.W. (1977) Selected Topics in Statistical Meteorology, AWS-TR-77-273, pp 7-1 to 7-45.
7. Chisholm, D.A., Lynch, R.H., Weyman, J.C., and Geisler, E.B. (1980) A Demonstration of the Modular Automated Weather System (MAWS), AFGL-TR-80-0087, AD A087070.
8. Weyman, J.C., and Lynch, R.H. (1981) A Digital Processing and Display System for the Rotating Beam Ceilometer (AN/GMQ-13), AFGL-TR-81-0027.
9. Geisler, E.B., and Chisholm, D.A. (1980) An Automated Cloud Observing System (ACOS) AFGL-TR-81-0082, AD A100266.
10. Duda, R.O., Mancuso, R.L., and Paskert, P.F. (1971) Analysis of Techniques for Describing the State of the Sky Through Automation, Report No. FAA-RD-71-52, prepared for Department of Transportation, Washington, D.C.
11. Miller, R.G., and Best, D.L. (1978) A model for converting probability forecasts to categorical forecasts, NOAA Tech. Memo, TDL Office Note 78-14.
12. Miller, R.G., and Best, D.L. (1979) A model for converting probability forecasts to categorical forecasts, Preprints, Sixth Conf. on Probability and Statistics in Atmospheric Sciences, Am. Meteorol. Soc., pp 98-102.
13. Brier, G.W. (1950) Verification of forecasts expressed in terms of probability, Mon. Wea. Rev., **78**:1-3.

Appendix A

Coefficient Sets for REEP Forecast Equations

1. REEP5 Prediction Equation Set

30-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0	0.009	0.124	0.679	0.180	0
1	0	0.003	-0.098	-0.538	0.641	5
2	0.727	0.183	-0.124	-0.679	-0.099	1
3	0.055	0.752	0.034	-0.664	-0.169	2
4	0	0.095	0.638	-0.584	-0.141	3
Total Reduction of Variance	0.539	0.518	0.478	0.357	0.564	

I. REEP5 Prediction Equation Set (Cont)

60-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0	0.188	0.632	0.108	0.066	0
1	0.005	-0.166	-0.600	0.085	0.682	5
2	0.626	0.025	-0.582	-0.078	0.015	1
3	0.081	0.441	-0.408	-0.080	-0.028	2
4	0	-0.167	-0.457	0.491	0.139	4
Total Reduction of Variance	0.381	0.321	0.304	0.240	0.433	

120-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0.009	0.247	0.506	0.130	0.102	0
1	0.547	0.005	-0.475	-0.100	0.029	1
2	-0.008	-0.204	-0.443	0.110	0.551	5
3	-0.006	-0.212	-0.249	0.300	0.172	4
4	0.090	0.234	-0.220	-0.062	-0.036	2
Total Reduction of Variance	0.293	0.183	0.162	0.099	0.278	

180-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0.026	0.232	0.451	0.125	0.159	0
1	0.418	0	-0.380	-0.024	-0.007	1
2	-0.024	-0.166	-0.356	0.176	0.376	5
3	0.086	0.201	-0.203	-0.038	-0.040	2
4	-0.022	-0.162	-0.188	0.229	0.149	4
Total Reduction of Variance	0.180	0.122	0.107	0.067	0.135	

2. **REEP.01 Prediction Equation Set**

30-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0	-0.001	0.092	0.695	0.207	0
1	0	0.008	-0.080	-0.546	0.626	5
2	0.523	-0.052	0.055	-0.590	0.070	1
3	-0.008	0.526	0.115	-0.585	-0.043	2
4	-0.002	0.046	0.585	-0.546	-0.077	3
5	0.257	0.242	-0.179	-0.107	-0.211	9
6	0.026	0.340	-0.100	-0.107	-0.157	8
7	-0.001	0.044	0.163	-0.074	-0.130	7
Total Reduction of Variance	0.594	0.559	0.512	0.358	0.575	

60-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0.001	0.180	0.695	0.096	0.022	0
1	0.012	-0.099	-0.549	0.041	0.599	5
2	0.406	-0.227	-0.323	0.024	0.126	1
3	0.012	0.247	-0.266	-0.013	0.025	2
4	0.006	-0.108	-0.420	0.454	0.070	4
5	0.278	0.273	-0.360	-0.100	-0.088	9
6	0.033	0.291	-0.248	-0.073	0	8
7	-0.006	-0.066	-0.125	0.063	0.139	6
Total Reduction of Variance	0.441	0.364	0.329	0.245	0.445	

2. **REEP.01 Prediction Equation Set (Cont)**

120-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0	0.179	0.384	0.218	0.207	0
1	0.322	0.290	-0.188	-0.133	-0.280	9
2	-0.001	-0.145	-0.337	0.034	0.460	5
3	0.048	0.345	-0.101	-0.158	-0.123	8
4	0.296	-0.213	-0.201	-0.052	0.169	1
5	0	-0.165	-0.163	0.240	0.098	4
6	0.014	0.100	0.176	-0.134	0.147	7
Total Reduction of Variance	0.358	0.202	0.185	0.112	0.294	

180-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0.063	0.250	0.322	0.216	0.138	0
1	0.277	-0.148	0.020	0.042	-0.189	9
2	-0.036	-0.150	-0.257	0.120	0.333	5
3	0.028	0.247	-0.139	-0.081	-0.046	2
4	-0.009	0.059	0.202	-0.100	-0.145	7
5	-0.035	-0.156	-0.115	0.183	0.132	4
6	-0.058	-0.102	-0.041	-0.041	0.247	13
7	0.187	0.129	-0.269	-0.127	0.088	1
8	-0.048	-0.076	0.054	-0.061	0.134	10
Total Reduction of Variance	0.258	0.143	0.144	0.080	0.195	

3. REEP11 Prediction Equation Sets

30-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0	0	0	0	0	0
1	0	0.009	-0.083	-0.550	0.623	5
2	0.523	-0.045	0.044	-0.584	0.062	1
3	-0.004	0.527	0.106	-0.582	-0.047	2
4	0.001	0.051	0.574	-0.542	-0.084	3
5	0.254	0.278	-0.106	0.598	-0.023	9
6	0.025	0.373	-0.028	0.599	0.031	8
7	-0.002	0.075	0.235	0.633	0.059	7
8	0	0.031	0.070	0.710	0.189	6
9	-0.008	-0.065	0.072	-0.050	0.051	13
10	0.009	-0.039	0.013	-0.002	0.019	12
11	-0.008	-0.031	0.030	-0.012	0.022	10
Total Reduction of Variance	0.596	0.561	0.515	0.364	0.576	

60-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0	0	0	0	0	0
1	0.008	-0.106	-0.540	0.031	0.607	5
2	0.401	-0.219	-0.331	0.018	0.130	1
3	0.013	0.241	-0.271	-0.010	0.028	2
4	0.002	-0.118	-0.408	0.442	0.082	4
5	0.264	0.477	0.341	-0.025	-0.058	9
6	0.022	0.491	0.448	0.008	0.031	8
7	-0.015	0.134	0.561	0.151	0.170	6
8	0.003	-0.074	0.069	-0.180	0.021	13
9	-0.011	0.200	0.694	0.085	0.033	7
10	0.025	-0.019	-0.028	0.049	-0.028	12
11	0.012	0.025	0.006	0.010	-0.053	11
Total Reduction of Variance	0.445	0.371	0.339	0.249	0.448	

3. **REEP11 Prediction Equation Sets (Cont)**

120-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0	0	0	0	0	0
1	0.301	0.198	-0.065	-0.128	-0.307	9
2	-0.025	-0.250	-0.212	0.046	0.441	5
3	0.033	0.272	-0.012	-0.150	-0.144	8
4	0.287	-0.220	-0.171	-0.058	0.162	1
5	-0.025	-0.273	-0.037	0.251	0.084	4
6	0.011	0.091	0.183	-0.138	-0.147	7
7	0.003	0.206	0.273	0.197	0.322	13
8	0.007	0.282	0.272	0.191	0.248	10
9	-0.020	-0.105	0.141	0.023	-0.028	3
10	0.040	0.333	0.291	0.185	0.151	11
11	0.042	0.294	0.223	0.236	0.206	12
Total Reduction of Variance	0.368	0.217	0.195	0.115	0.307	

180-min Forecast						
Order(i)	P_1	P_2	P_3	P_4	P_5	Predictor X_i
0	0	0	0	0	0	0
1	0.361	0.167	0.400	0.190	-0.118	9
2	-0.280	-0.121	-0.233	0.082	0.299	5
3	0.008	0.177	-0.201	-0.006	0.022	2
4	0.056	0.302	0.519	0.127	-0.004	7
5	-0.027	-0.130	-0.094	0.150	0.102	4
6	-0.060	-0.092	-0.033	-0.055	0.240	13
7	0.166	0.058	-0.332	-0.051	0.159	1
8	-0.050	-0.069	0.060	-0.072	0.131	10
9	0.096	0.350	0.409	0.111	0.034	8
10	0.055	0.209	0.287	0.269	0.180	6
11	-0.077	0.026	0.024	-0.032	-0.010	11
Total Reduction of Variance	0.260	0.150	0.149	0.087	0.200	

Appendix B

Transition Matrix

The following transition matrix was used in the Equivalent Markov prediction technique and can be interpreted in the following way:

$$\underset{\sim}{P} \quad \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{15} \\ P_{21} & P_{22} & \cdots & P_{25} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ P_{51} & P_{52} & \cdots & P_{55} \end{bmatrix}$$

where

$$P_{ij} = P[\text{CEILING}_t = i \mid \text{CEILING}_0 = j] \text{ for } i, j, = 1, 2, \ldots, 5$$

That is, P_{ij} is the transition or conditional probability that the ceiling category at time t will be in state i given the ceiling category at time 0 was in state j.

30-min One-Step Transition Matrix

$$\underset{\sim}{P} = \begin{bmatrix} 0.727 & 0.192 & 0.000 & 0.000 & 0.081 \\ 0.055 & 0.760 & 0.158 & 0.015 & 0.011 \\ 0.000 & 0.104 & 0.762 & 0.095 & 0.039 \\ 0.000 & 0.009 & 0.125 & 0.684 & 0.181 \\ 0.000 & 0.011 & 0.026 & 0.141 & 0.821 \end{bmatrix}$$

Appendix C

Coefficient Set for REEP Equation Categorical Forecasts

The following set of REEP equations was used in the iterative and unit bias thresholding models (see Section 4). Each forecast equation generates a probability estimate (P) of a ceiling less than 1000 ft. See Table 4 for a summary of predictor limits. For example, what is the probability a ceiling less than 1000 ft will occur in 30 min given the observation matrix $\underset{\sim}{Q}$ = (1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0,)?

$$P = P(\text{CEILING} < 1000 \text{ ft}) = 0.425 + 0.523 = 0.948 .$$

Also,

$$P(\text{CEILING} \geq 1000 \text{ ft or no ceiling}) = 1.0 - P = 0.052 .$$

PRECEDING PAGE BLANK-NOT FILMED

1. REEP Coefficients Estimate for Ceiling Less Than 1000 ft

Order(i)	Forecast Length (min) 30 P	30 Predictor X(i)	60 P	60 Predictor X(i)
0	0.425	0	0.942	0
1	-0.324	6	-0.282	6
2	0.626	3	-0.620	5
3	0.629	2	-0.506	4
4	0.523	1	-0.078	7
5	-0.073	5	0.044	11
6	-0.118	7	-0.131	1
7	-0.055	8	0.122	9
8	-0.017	12	-0.020	12
9	-0.010	10	0.018	3
Total Reduction of Variance	0.850		0.778	

Order(i)	Forecast Length (min) 120 P	120 Predictor X(I)	180 P	180 Predictor X(i)
0	0.857	0	0.875	0
1	-0.284	6	-0.325	6
2	-0.500	5	-0.379	5
3	-0.348	4	-0.249	4
4	0.104	11	-0.185	13
5	-0.078	13	-0.059	10
6	0.138	9	0.041	11
7	-0.106	1	-0.094	1
8	------	--	-0.031	8
9	------	--	0.038	9
Total Reduction of Variance	0.679		0.594	

DATE
FILMED

8